Age Prediction in Low-Resolution Images: A Comparative Study of Deep Learning and Super-Resolution Models

Jonathan Adrian[1*] and Gede Putra Kusuma [2]
[1-2]Computer Science Department, BINUS Graduate Program –
Master of Computer Science, Bina Nusantara University,
Jakarta, Indonesia 11480
Email: [1]jonathan.adrian001@binus.ac.id, [2]inegara@binus.edu

Abstract *– This study aims to enhance age prediction accuracy from low-resolution (LR) facial images by evaluating and comparing the effectiveness of deep learning models and super-resolution techniques. The research addresses the challenge of accurate age estimation in real-world scenarios where LR images are prevalent, which often degrade the performance of traditional models. Three super-resolution models—SRCNN, RCAN, and EDSR—were employed to reconstruct high-resolution (HR) images from LR inputs, while three age prediction models—CORAL CNN, GRA_Net, and Attention VGG-16—were tested for their ability to estimate age from both high-resolution and super-resolution (SR) images. The experiments were conducted using the UTKFace dataset, with performance evaluated using metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), R² score, and Structural Similarity Index (SSIM). Results demonstrated that the CORAL CNN model achieved the best age prediction performance on HR images, with an MAE of 6.2319, an MSE of 73.8212, and an R² score of 0.8181. Among the super-resolution models, RCAN outperformed others, achieving an SSIM of 0.9129 and an MSE of 0.0009. When combined with CORAL CNN, RCAN further improved age prediction accuracy on SR images, outperforming traditional bicubic interpolation. This study demonstrate the potential of combining super-resolution techniques with deep learning models to address resolution limitations in age prediction tasks.*

Keywords: age prediction, super-resolution, deep learning, facial analysis, low-resolution image

I. INTRODUCTION

Age prediction from facial images is a critical area of research with significant applications in security, marketing, and personalized services. However, in real-world scenarios, face images are often of low resolution, posing substantial challenges for accurate age estimation. Traditional models, particularly Convolutional Neural Networks (CNNs), have demonstrated exceptional performance in tasks such as face recognition and age prediction when applied to high-resolution (HR) images [1, 2]. CNNs excel at extracting relevant features necessary for classification tasks. However, their performance deteriorates significantly on low-resolution (LR) images, as key facial features—such as skin texture, wrinkles, and fine lines—become less discernible [3, 4].

While age prediction has been explored in various domains, including demographic prediction and human-computer interaction [5, 6], most existing models remain heavily reliant on high-quality images. This dependence becomes problematic in real-world environments, such as video surveillance or low-quality mobile devices, where LR images are prevalent. Additionally, datasets for training face recognition models often focus on variations like age, gender, and pose but generally overlook resolution variability. This oversight results in a significant performance gap when these models are tested on LR images [3].

A promising solution to this challenge is the use of super-resolution techniques. These models enhance image quality by reconstructing HR images from LR inputs, enabling deep learning models to extract more detailed features for age estimation [7]. By improving image quality prior to processing, super-resolution techniques offer a viable approach to enhancing age prediction accuracy, particularly in low-resolution scenarios. This has spurred growing interest in integrating super-resolution models with age estimation methods [5].

Despite these advancements, challenges persist, particularly the lack of specialized training datasets that account for LR images. Most existing datasets emphasize factors such as age, gender, and pose, but resolution differences are often neglected. Consequently, models trained on such datasets perform well on HR images but struggle to maintain accuracy when applied to LR face images, which are more commonly encountered in real-world applications [3].

This study aims to evaluate and compare the effectiveness of deep learning models and super-resolution techniques in improving age prediction accuracy from LR facial images. The age prediction models used in this research include CORAL CNN [18], GRA_Net [9], and Attention VGG-16 [31], while the super-resolution models used are SRCNN [27, 28], RCAN [29], and EDSR [30]. By assessing these methods, the research seeks to identify more reliable approaches that can address resolution limitations, ultimately enhancing the robustness of age estimation in practical applications. Given the increasing demand for reliable face recognition systems in fields like healthcare and public security [8, 9], advancing age prediction technologies is essential for improving system reliability and expanding their applicability.

II. RELATED WORKS

2.1 Super-Resolution

Several studies have focused on enhancing facial images in challenging conditions, such as low resolution and partial occlusion. Cai et al. (2019) proposed FCSR-GAN, which combines face completion and super-resolution to improve partially occluded facial images and enhance the quality of unobstructed facial images [10]. Zhang and Ling (2020) introduced SPGAN, which addresses extremely low-resolution faces (16×16 pixels) by considering facial identity and improves face recognition accuracy [11]. These approaches demonstrate the effectiveness of generative adversarial networks (GANs) in handling low-resolution and occluded facial images.

Chen et al. (2020) presented SPARNet, a spatial attention residual network that enhances facial recovery from low-resolution (LR) images and generates high-resolution outputs [12]. Muqeet et al. (2019) proposed HRAN, optimizing multiscale feature extraction through binarization feature fusion, improving image quality and computational efficiency [13].

Massoli et al. (2020) addressed cross-resolution face recognition challenges, demonstrating that their method enhances recognition on low-resolution images while boosting the performance of super-resolution models [3]. Chen et al. (2020) also proposed an identity-based method to recover identity information from low-resolution faces [14]. Additionally, Cheng et al. (2019) introduced CSRI, which integrates super-resolution and identity preservation to improve face recognition performance on low-resolution images [15].

Vo et al. (2020) proposed PSR to enhance facial expression recognition in challenging in-the-wild images, considering factors such as low resolution, pose, and orientation [16].

Recent studies have begun to explore the application of super-resolution techniques to age prediction tasks. For example, Nam et al. (2020) proposed a deep convolutional neural network (CNN)-based age estimation method that reconstructs low-resolution facial images as high-resolution (HR) images using a conditional generative adversarial network (GAN). This hybrid approach demonstrates that preserving age-related features during upscaling significantly improves age prediction accuracy on low-resolution images. The method was validated on two open databases (PAL and MORPH), achieving higher accuracy in both high-resolution reconstruction and age estimation compared to state-of-the-art methods [17].

2.2 Age Prediction

Cao et al (2020) propose the COnsistent RAnk Logits (CORAL) framework, an ordinal regression method that ensures rank-monotonicity and consistent confidence scores. This method demonstrate a substantial reduction in prediction errors compared to traditional ordinal regression networks [18]. Yudin et al. (2019) used ResNet and Xception on an imbalanced dataset for age classification (101 classes) and gender classification (binary), focusing on computational efficiency using NVIDIA CUDA [19].

Sikder et al. (2022) introduced a multitask method to recognize emotions, gender, and age simultaneously, combining M-CNN and M-SVM, achieving high accuracy on a custom dataset [20]. Georgescu et al. (2021) addressed the challenges of age and gender recognition on occluded faces by using knowledge distillation to improve CNN performance, showing significant results on VGG and ResNet-50 models [21].

Garain et al. (2021) proposed GRA_Net with a gated attention mechanism for age and gender prediction, combining classification and regression to improve accuracy on datasets such as FG-Net and UTKFace [9]. Terhörst et al. (2019) introduced a reliability measure for age and gender prediction models using dropout-reduced neural networks, demonstrating a correlation between reliability and prediction performance [22].

Sheoran et al. (2021) found that pre-trained models such as VGG16 and ResNet50 outperformed CNNs trained from scratch, with transfer learning yielding competitive results even with simple algorithms [23].

III. METHODOLOGY

3.1 Dataset

The UTKFace dataset is a comprehensive collection of facial images designed for age, gender, and ethnicity analysis. It consists of 23,689 uniformly sized images (200 × 200 pixels) [21], representing individuals across a wide age range from 0 to 116 years. The dataset is richly annotated with labels for age, gender, and ethnicity. Additionally, the images exhibit significant variations in factors such as pose, facial expressions, lighting conditions, occlusions, and resolution, ensuring robustness for training and evaluation purposes.

For experimental setup, the dataset is partitioned into three subsets: 18,951 images (80%) are allocated for training, while 2,368 images (10%) are used for validation and another 2,368 images (10%) for testing. This division ensures a balanced evaluation of model performance across different stages of development. Figure 1 provides sample images from the dataset.

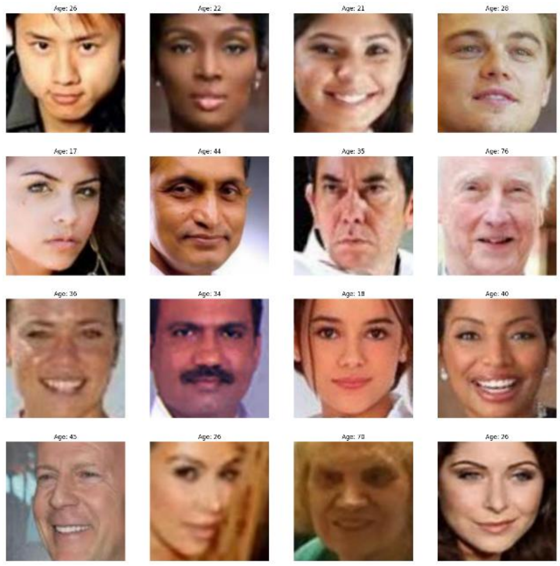

Figure 1 Sample images from the UTKFace dataset

3.2 Preprocessing

The facial images in the dataset were preprocessed to ensure consistency and compatibility with deep learning models. All images were resized to a uniform resolution: high-resolution (HR) images were scaled to 128 × 128 pixels, while low-resolution (LR) images were downsampled to 32 × 32 pixels. After resizing, pixel values were normalized to the range of 0 to 1 by dividing each intensity value by 255.

To further enhance the robustness of the model and prevent overfitting, data augmentation techniques are applied only to the training data using the ImageDataGenerator function. The augmentation settings included a rotation range of 15 degrees, horizontal and vertical shifts of 10% of the image dimensions, a shear range of 0.1, a zoom range of 0.1, horizontal flipping enabled, and a fill mode set to 'nearest' for handling boundary pixels [24]. These augmentations were designed to increase the diversity of the training data, thereby improving the model's generalization capability.

3.3 Super-Resolution

Face super-resolution, or face hallucination, is a specialized technique aimed at generating high-resolution (HR) face images from low-resolution (LR) inputs [25, 26]. This section discusses three deep learning models used for this task: Super-Resolution Neural Network (SRCNN), Residual Channel Attention Network (RCAN), and Enhanced Deep Super-Resolution (EDSR). These models were selected for their effectiveness in enhancing image quality and reconstructing features, textures, and facial structures that are not clearly visible in low-resolution images, which are critical for applications such as surveillance and face recognition.

The first model, Super-resolution Neural Network (SRCNN), is a pioneering convolutional neural network (CNN) designed for super-resolution tasks [27]. It uses convolutional layers to learn an end-to-end mapping between low and high-resolution images through three key operations: patch extraction and representation, non-linear mapping, and reconstruction. These operations are implemented using convolutional layers that extract features, apply non-linear transformations, and reconstruct the high-resolution image [28].

The second model, Residual Channel Attention Network (RCAN), is a deep learning model for single-image super-resolution (SISR). It leverages a Residual-in-Residual (RIR) architecture and a Channel Attention (CA) mechanism to enhance image quality. The model consists of four main components: shallow feature extraction, deep feature extraction using RIR (including Residual Groups (RGs) and Residual Channel Attention Blocks (RCABs)), an upscaling module, and a final reconstruction layer. The CA mechanism dynamically adjusts the importance of each feature channel, enabling the model to focus on more informative features for better reconstruction [29].

The third model, Enhanced Deep Super-Resolution (EDSR), improves upon the SRResNet architecture by removing batch normalization layers, which reduces GPU memory usage by 40% and allows for a deeper network with 32 residual blocks and 256 feature maps per layer. To stabilize training, EDSR employs residual scaling (0.1) and a pretraining strategy, where a model trained for ×2 upscaling initializes training for ×3 and ×4 upscaling [30].

3.4 Model for Age Prediction

Age prediction based on facial images is a task in computer vision that aims to predict a person's age by analyzing facial characteristics. Features such as skin texture, wrinkles, and facial structure provide important information for age prediction. These characteristics change over time due to the natural aging process [32, 33, 34]. This section discusses three deep learning models for age prediction: Consistent Rank Logits (CORAL), Gated Residual Attention Network (GRA_Net), and Attention VGG-16.

The first model, Consistent Rank Logits (CORAL), addresses the challenge of classifier inconsistency in ordinal regression by transforming the problem into a series of binary classification tasks. Each task predicts whether a given rank exceeds a predefined threshold. A shared weight across all binary classifiers ensures rank-monotonicity. The model uses a Convolutional Neural Network (CNN), and the final rank prediction is based on the consistency of the binary classifier outputs. The loss function is a weighted cross-entropy, and the framework guarantees consistent rank predictions by ensuring non-increasing bias units in the output layer [18].

The second model, Gated Residual Attention Network (GRA_Net), integrates attention mechanisms, residual learning, and gating to enhance feature selection and improve robustness, particularly in noisy data environments. The architecture comprises multiple Attention Blocks, each consisting of a Trunk Branch for feature processing and a Mask Branch that generates a mask to modulate the output features. A gating mechanism regulates the influence of the mask, preserving the advantages of residual learning. By stacking Attention Blocks, GRA_Net progressively refines feature representations, achieving improved performance with increased depth. For regression tasks, the final output layer produces continuous predictions, making it suitable for age estimation [9].

The third model, Attention VGG-16, enhance age and gender recognition through a three-module architecture. The first module, an Attention CNN, generates an attention map to identify important image regions. The second module, a Patch CNN, processes high-resolution patches guided by the attention map, reducing spatial dimensions via Global Average Pooling. The third module, an MLP classifier, combines features from both CNNs to produce the final classification. Attention is implemented using either a weighted sum (soft attention) or an element-wise product (hard attention) for efficiency [31].

3.5 Evaluation Process

After training, the model's performance is evaluated using test data. Performance metrics are used to assess the effectiveness of each model.

3.5.1 Evaluation of Age Prediction Based on HR Images

The performance of the age prediction model using HR (High-Resolution) images is evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), and R^2 score. MAE measures the average absolute difference between the predicted and actual ages, while MSE calculates the average squared difference. The R^2 score measures how well the model explains the variance in actual ages, with a value closer to 1 indicating a better fit to the data. Lower MAE and MSE values, along with a higher R^2 score, reflect greater accuracy in age prediction.

3.5.2 Evaluation of Super-Resolution Model

The super-resolution model's performance is evaluated using Structural Similarity Index (SSIM) and Mean Squared Error (MSE). SSIM measures the similarity between the reconstructed image and the original HR image, with higher SSIM values indicating greater similarity. In contrast, MSE measures the average squared difference between the reconstructed and HR images, with lower values indicating better performance.

3.5.3 Evaluation of Age Prediction Based on SR Images

The performance of the age prediction model using SR (Super-Resolution) images is evaluated using MAE, MSE, and R^2 score. For comparison, the model's performance is compared against Bicubic interpolation, a traditional method for image upscaling. The goal is for the SR-based model's performance to fall between that of the HR-based model and the Bicubic interpolation method, demonstrating its effectiveness in bridging the gap between low-resolution and high-resolution image analysis.

IV. RESULTS AND DISCUSSION

This section presents the results of the experiments conducted for age prediction and super-resolution tasks using the UTKFace dataset. The performance of the models is evaluated using standard metrics, and the findings are discussed in detail. The results are compared across different models and configurations to identify the most effective approaches.

4.1 Evaluation Results of Age Prediction based on HR Images

The age prediction model was trained for 100 epochs using the UTKFace dataset with HR images. The model performance was evaluated using test data, and the results are presented in Table 1. The evaluation metrics used include Mean Absolute Error (MAE), Mean Squared Error (MSE), and R^2 score, which measure the accuracy and consistency of the predictions.

Table 1 Performance Result based on HR Images on Test Data

Model	**MAE**	**MSE**	**R^2 Score**
CORAL CNN	**6.2319**	**73.8212**	**0.8181**
GRA_Net	6.8758	90.6539	0.7766
Attention VGG-16	7.4081	108.2656	0.7333

From Table 1, the CORAL CNN model outperformed the other models, achieving the lowest MAE of 6.2319 and MSE of 73.8212 values, as well as the highest R^2 score of 0.8181. GRA_Net showed the second-best performance, while Attention VGG-16 lagged behind.

4.2 Evaluation Results of Super-Resolution Model

The super-resolution models were trained to reconstruct LR images (32 × 32 pixels) into SR images (128 × 128 pixels), with HR images used as the ground truth. The models were evaluated using Structural Similarity Index (SSIM) and Mean Squared Error (MSE), which measure the quality of the reconstructed images.

Table 2 Super Resolution Performance Result on Test Data

Model	**SSIM**	**MSE**
SRCNN	0.8975	0.0017
RCAN	**0.9129**	**0.0009**
EDSR	0.8923	0.0012

From Table 2, RCAN achieved the best performance, with the highest SSIM of 0.9129 and the lowest MSE of 0.0009. This can be attributed to its Residual-in-Residual (RIR) architecture and Channel Attention (CA) mechanism, which enhance feature extraction and reconstruction. SRCNN showed a higher SSIM than EDSR but a higher MSE, indicating a trade-off between structural similarity and pixel-level accuracy. EDSR, despite its deeper architecture, underperformed slightly, possibly due to its removal of batch normalization layers, which may have affected stability during training.

4.3 Evaluation Results of Age Prediction based on SR Images

The age prediction model was combined with the super-resolution models to evaluate its performance on SR images. Based on Table 1, CORAL CNN, which showed the best performance on HR images, was selected for this task. The results are compared with bicubic interpolation, a traditional upscaling method.

Table 3 Performance Result based on SR Images on Test Data

Model	**MAE**	**MSE**	**R^2 Score**
CORAL CNN + SRCNN	6.3550	76.6014	0.8032
CORAL CNN + RCAN	**6.2617**	**75.3740**	**0.8063**
CORAL CNN + EDSR	6.9240	90.4205	0.7677
CORAL CNN + Bicubic Interpolation	6.4629	80.6361	0.8013

From Table 3, the combination of CORAL CNN with RCAN achieved the best performance, with an MAE of 6.2617, MSE of 75.3740, and R^2 score of 0.8063. This demonstrates the effectiveness of RCAN in reconstructing high-quality SR images, and improving the accuracy of age prediction. All super-resolution models outperformed bicubic interpolation, showing the advantages of deep learning-based approaches. However, the EDSR model underperformed, possibly due to its lower SSIM and higher MSE values compared to RCAN and SRCNN.

V. CONCLUSION AND FUTURE WORK

This study analyzed the effectiveness of deep learning models and super-resolution techniques in improving age prediction accuracy from low-resolution (LR) facial images. The evaluation shows that the CORAL CNN model outperformed other age prediction models when applied to high-resolution (HR) images, achieving the lowest MAE of 6.2319, the lowest MSE of 73.8212 and the highest R^2 score of 0.8181. Among the super-resolution models, RCAN demonstrated the best performance, with the highest SSIM of 0.9129 and the lowest MSE of 0.0009. When combined with CORAL CNN, RCAN further improved the accuracy of age prediction in super-resolution (SR) images, achieving an MAE of 6.2617, an MSE of 75.3740, and an R^2 score of 0.8063, outperforming traditional bicubic interpolation. These results highlight the potential of integrating super-resolution techniques with deep learning models to bridge the gap between LR and HR image analysis, thereby increasing the robustness of age prediction systems in real-world applications.

Although the results are promising, several challenges remain. The lack of specialized datasets containing LR images with varying resolution levels limits the generalizability of the models. Exploring hybrid architectures that combine the strengths of different super-resolution and age prediction techniques could lead to further improvements. Additionally, extending this research to related tasks, such as gender prediction, and evaluating the models in more challenging real-world scenarios, such as noisy images, would provide a more comprehensive understanding of its applicability. Overcoming these challenges will pave the way for more robust and efficient age prediction systems across various practical situations.

REFERENCES

[1] S. Chattopadhyay, A. Dey, Pawan Kumar Singh, D. Oliva, E. Cuevas, and R. Sarkar, "MTRRE-Net: A deep learning model for detection of breast cancer from histopathological images," *Computers in Biology and Medicine*, vol. 150, p. 106155, Nov. 2022, doi: 10.1016/j.compbiomed.2022.106155.

[2] K. Khan, M. Attique, R. U. Khan, I. Syed, and T.-S. Chung, "A Multi-Task Framework for Facial Attributes Classification through End-to-End Face Parsing and Deep Convolutional Neural Networks," *Sensors*, vol. 20, no. 2, p. 328, Jan. 2020, doi: 10.3390/s20020328.

[3] F. V. Massoli, G. Amato, and F. Falchi, "Cross-resolution learning for Face Recognition," *Image and Vision Computing*, vol. 99, p. 103927, Jul. 2020, doi: 10.1016/j.imavis.2020.103927.

[4] M. Rouhsedaghat, Y. Wang, X. Ge, S. Hu, S. You, and J. Kuo, "FaceHop: A Light-Weight Low-Resolution Face Gender Classification Method," in *Pattern Recognition. ICPR International Workshops and Challenges*, Milan, Italy, Jan. 10–15, 2021, pp. 169–183.

[5] M. Fayyaz, M. Yasmin, M. Sharif, and M. Raza, "J-LDFR: joint low-level and deep neural network feature representations for pedestrian gender classification," *Neural Computing and Applications*, May 2020, doi: 10.1007/s00521-020-05015-1.

[6] H. A. Sánchez-Hevia, R. Gil-Pita, M. Utrilla-Manso, and M. Rosa-Zurera, "Age group classification and gender recognition from speech with temporal convolutional neural networks," *Multimedia Tools and Applications*, vol. 81, pp. 3535–3552, Jan. 2022.

[7] H. Yousif and Zahraa Al-Milaji, "Fault detection from PV images using hybrid deep learning model," *Solar energy*, vol. 267, p. 112207, Jan. 2024, doi: 10.1016/j.solener.2023.112207.

[8] O. Agbo-Ajala and S. Viriri, "Deep learning approach for facial age classification: a survey of the state-of-the-art," *Artificial Intelligence Review*, vol. 54, pp. 179–213, Jun. 2020, doi: 10.1007/s10462-020-09855-0.

[9] A. Garain, B. Ray, P. K. Singh, A. Ahmadian, N. Senu, and R. Sarkar, "GRA_Net: A Deep Learning Model for Classification of Age and Gender From Facial Images," *IEEE Access*, vol. 9, pp. 85672–85689, Jun. 2021, doi: 10.1109/access.2021.3085971.

[10] J. Cai, H. Hu, S. Shan, and X. Chen, "FCSR-GAN: End-to-end Learning for Joint Face Completion and Super-resolution," in *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition*, Lille, France, May 14-18, 2019, doi: 10.1109/fg.2019.8756607.

[11] M. Zhang and Q. Ling, "Supervised Pixel-Wise GAN for Face Super-Resolution," *IEEE Transactions on Multimedia*, vol. 23, pp. 1938–1950, Jul. 2020, doi: 10.1109/tmm.2020.3006414.

[12] C. Chen, D. Gong, H. Wang, Z. Li, and K.-Y. K. Wong, "Learning Spatial Attention for Face Super-Resolution," *IEEE Transactions on Image Processing*, vol. 30, pp. 1219–1231, Dec. 2020, doi: 10.1109/TIP.2020.3043093.

[13] A. Muqeet, M. T. B. Iqbal, and S.-H. Bae, "HRAN: Hybrid Residual Attention Network for Single Image Super-Resolution," *IEEE Access*, vol. 7, pp. 137020–137029, Sep. 2019, doi: 10.1109/access.2019.2942346.

[14] J. Chen, J. Chen, Z. Wang, C. Liang, and C.-W. Lin, "Identity-Aware Face Super-Resolution for Low-Resolution Face Recognition," *IEEE Signal Processing Letters*, vol. 27, pp. 645–649, Apr. 2020, doi: 10.1109/lsp.2020.2986942.

[15] Z. Cheng, X. Zhu, and S. Gong, "Low-Resolution Face Recognition," in *Computer Vision – ACCV 2018*, Perth, Australia, Dec. 2–6, 2018, pp. 605–621.

[16] T. Vo, G. Lee, H. Yang, and S. Kim, "Pyramid With Super Resolution for In-the-Wild Facial Expression Recognition," *IEEE Access*, vol. 8, pp. 131988–132001, Jul. 2020, doi: 10.1109/access.2020.3010018.

[17] S. H. Nam, Y. H. Kim, N. Q. Truong, J. Choi, and K. R. Park, "Age Estimation by Super-Resolution Reconstruction Based on Adversarial Networks," *IEEE Access*, vol. 8, pp. 17103–17120, Jan. 2020, doi: 10.1109/access.2020.2967800.

[18] W. Cao, V. Mirjalili, and S. Raschka, "Rank consistent ordinal regression for neural networks with application to age estimation," *Pattern Recognition Letters*, vol. 140, pp. 325–331, Dec. 2020, doi: 10.1016/j.patrec.2020.11.008.

[19] D. Yudin, M. Shchendrygin, and A. Dolzhenko, "Age and Gender Recognition on Imbalanced Dataset of Face Images with Deep Learning," in *Proceedings of the Fourth International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'19)*, Ostrava–Prague, Czech Republic, Dec. 2–7, 2019, pp. 30–40.

[20] J. Sikder, N. Datta, Sajib Tripura, and U. K. Das, "Emotion, Age and Gender Recognition using SURF, BRISK, M-SVM and Modified CNN," in *2022 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, Prague, Czech Republic, Jul. 20–22, 2022, doi: 10.1109/icecet55527.2022.9872771.

[21] M.-I. Georgescu, G.-E. Duţă, and R. T. Ionescu, "Teacher–student training and triplet loss to reduce the effect of drastic face occlusion," *Machine Vision and Applications*, vol. 33, no. 12, Dec. 2021.

[22] P. Terhörst *et al.*, "Reliable Age and Gender Estimation from Face Images: Stating the Confidence of Model Predictions," in *2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, Tampa, FL, USA, Sep. 23–26, 2019, doi: 10.1109/btas46853.2019.9185975.

[23] V. Sheoran, S. Joshi, and T. R. Bhayani, "Age and Gender Prediction Using Deep CNNs and Transfer Learning," in *Computer Vision and Image Processing*, Prayagraj, India, Dec. 4–6, 2020, pp. 293–304.

[24] P. Dey, T. Mahmud, M. S. Chowdhury, M. S. Hossain, and K. Andersson, "Human Age and Gender Prediction from Facial Images Using Deep Learning Methods," *Procedia Computer Science*, vol. 238, pp. 314–321, 2024, doi: 10.1016/j.procs.2024.06.030.

[25] M. Li, Z. Zhang, J. Yu, and C. W. Chen, "Learning Face Image Super-Resolution Through Facial Semantic Attribute Transformation and Self-Attentive Structure Enhancement," *IEEE Transactions on Multimedia*, vol. 23, pp. 468–483, Apr. 2020, doi: 10.1109/tmm.2020.2984092.

[26] Z.-S. Liu, W.-C. Siu, and Y.-L. Chan, "Features Guided Face Super-Resolution via Hybrid Model of Deep Learning and Random Forests," *IEEE Transactions on Image Processing*, vol. 30, pp. 4157–4170, 2021, doi: 10.1109/tip.2021.3069554.

[27] Y. K. Ooi and H. Ibrahim, "Deep Learning Algorithms for Single Image Super-Resolution: A Systematic Review," *Electronics*, vol. 10, no. 7, p. 867, Apr. 2021, doi: 10.3390/electronics10070867.

[28] C. Dong, C. C. Loy, K. He, and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295–307, Jun. 2015, doi: 10.1109/tpami.2015.2439281.

[29] Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image Super-Resolution Using Very Deep Residual Channel Attention Networks," in *Computer Vision – ECCV 2018*, Munich, Germany, Sep. 8–14, 2018, pp. 294–310, 2018.

[30] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced Deep Residual Networks for Single Image Super-Resolution," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, HI, USA, Jul. 21–26, 2017, doi: 10.1109/cvprw.2017.151.

[31] P. Rodríguez, G. Cucurull, J. M. Gonfaus, F. X. Roca, and J. Gonzàlez, "Age and gender recognition in the wild with deep attention," *Pattern Recognition*, vol. 72, pp. 563–571, Dec. 2017, doi: 10.1016/j.patcog.2017.06.028.

[32] Y. Lin, J. Shen, Y. Wang, and Maja Pantic, "FP-Age: Leveraging Face Parsing Attention for Facial Age Estimation in the Wild," *IEEE Transactions on Image Processing*, pp. 1–1, Mar. 2022, doi: 10.1109/tip.2022.3155944.

[33] S. Hiba and Y. Keller, "Hierarchical Attention-Based Age Estimation and Bias Analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 12, pp. 14682–14692, Sep. 2023, doi: 10.1109/tpami.2023.3319472.

[34] H. Zhu, Y. Zhang, G. Li, J. Zhang, and H. Shan, "Ordinal distribution regression for gait-based age estimation," *Science China Information Sciences*, vol. 63, Jan. 2020.